Universidade Federal
do Espírito Santo

CENTRO TECNOLÓGICO (CT)
DEPARTAMENTO DE INFORMÁTICA (DI)
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA (PPGI)

Laboratório de Redes – 2022/1 – DI/PPGI

Prof. Rodolfo da Silva Villaça – rodolfo.villaca@ufes.br

Objetivos Gerais:

- Aplicar técnicas de medição e monitoramento para aquisição de métricas de desempenho e falhas em redes de computadores;
- Aplicar técnicas de aprendizado de máquinas para, a partir dos dados de monitoramento, inferir conhecimento a respeito do funcionamento das redes de computadores.

Objetivos Específicos:

- Trabalho I: Aplicar técnicas de aprendizado federado (Federated Learning) para predição de métricas de QoS (Qualidade de Serviço) em aplicações de transmissão de vídeo;
- Trabalho II: Aplicar técnicas de aprendizado de máquinas para detecção de anomalias em redes de computadores a partir da medição de tráfego por meio de estruturas de dados probabilísticas (sketches),
- Trabalho III: Aplicar técnicas de aprendizado de máquinas para detecção e prevenção de intrusão em redes e sistemas computacionais por meio da análise de tráfego, fluxo e pacotes.

A partir destes objetivos, seguem alguns requisitos, sugestões de metodologia de execução e bibliografias para a execução dos Trabalhos I, II e III.

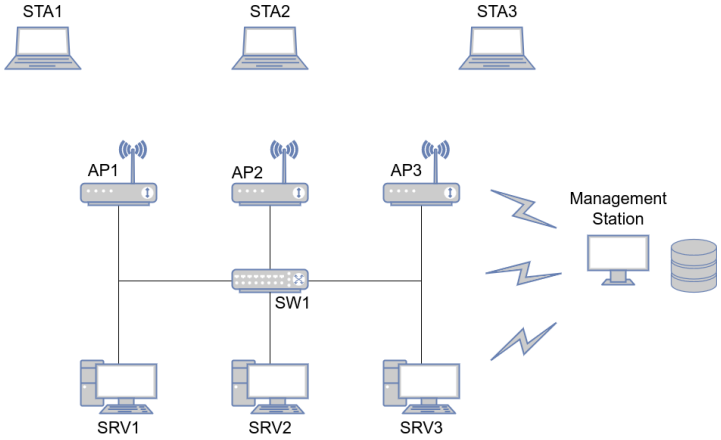
Universidade Federal
do Espírito Santo

CENTRO TECNOLÓGICO (CT)
DEPARTAMENTO DE INFORMÁTICA (DI)
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA (PPGI)

Trabalho I – Aprendizado Federado + QoS

Integrantes:
- Lucas Miguel Tassis;
- Eduardo Montagner de Moraes Sarmento.

Sugestão:



Requisitos:

- Servidores SRV1, SRV2, SRV3: devem executar um serviço de *streaming* de vídeo e possuir seus indicadores de desempenho de sistemas monitorados continuamente (utilização de CPU, ocupação de memória, utilização de disco, tráfego nas interfaces de rede e etc…);
- SW1, AP1, AP2, AP3: deve possuir seus indicadores de desempenho monitorados continuamente durante todo o período de experimentação. Sugere-se o uso do sFlow-RT[1] para essa tarefa;
- STA1, STA2, STA3: deve executar um cliente de reprodução de *streaming* de vídeos (tipo VLC, por exemplo) e ativar o *log* do serviço para obtenção de métricas de qualidade locais, que servirão como os *labels* (*targets*) para o aprendizado federado;
- Importante: o período de amostragem nos servidores, *hosts* móveis, *switches* e pontos de acesso deve ser compatível para facilitar o trabalho durante o treinamento dos preditores;

1 https://sflow-rt.com/index.php

Metodologia:

- No aprendizado federado, a entrada para treinamento do treinador i (i = 1, 2, 3) seria X <sw_i, ap_i, srv_i> e Y<sta_i>. Sugere-se avaliar a inclusão de dados de todos os APs no treinamento como *baseline* de avaliação X <sw_i, ap_1, ap_2, ap_3, srv_i>, Y <sta_i>;
- Os clientes móveis usam sempre o mesmo servidor durante todo o experimento: STA1 ↔ SRV1, STA2 ↔ SRV2, STA3 ↔ SRV3;
- Os clientes móveis devem mover-se durante todo o experimento. O padrão de movimentação pode ser definido pelo grupo.

Bibliografia:

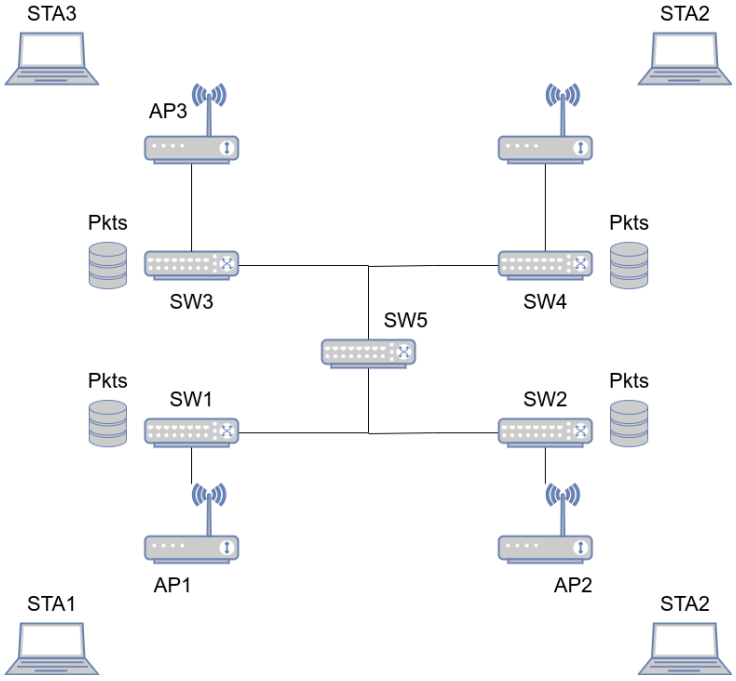
- [1] Stadler, R., Pasquini, R. & Fodor, V. Learning from Network Device Statistics. J Netw Syst Manage 25, 672–698 (2017).
- [2] R. Ul Mustafa, M. T. Islam, C. Rothenberg, S. Ferlin, D. Raca and J. J. Quinlan. DASH QoE Performance Evaluation Framework with 5G Datasets. 2020 16th International Conference on Network and Service Management (CNSM), Izmir, Turkey, 2020, pp. 1-6, doi: 10.23919/CNSM50824.2020.9269111.
- Integration between Mininet-WiFi and sflow-rt
- Providing Mobility with Mininet-WiFi

Trabalho II – Anomalias de Tráfego + Sketches

Integrantes:

- Athus Assunção Cavalini;
- Thiago da Silva Meireles de Souza;
- Iago de Sousa Cerqueira.

Sugestão:



Requisitos:

- SW1, SW2, SW3, SW4, SW5, SW6: deve-se coletar todos os pacotes que trafegam nestes dispositivos para posterior processamento. Sugere-se o uso das ferramentas *wireshark* ou *tcpdump* para essa tarefa;
- STA1, STA2, STA3, STA4: devem mover-se durante todo o experimento e realizar a troca de tráfego entre si. Sugere-se o uso da ferramenta iperf para a geração de tráfego entre as estações móveis.

Metodologia:

- O padrão de movimentação das estações móveis e de geração de tráfego pode ser definido pelo grupo. Por exemplo, pode-se definir que o padrão de movimentação é aleatório e o tráfego é uniformemente distribuído em tempo/volume entre as estações; pode-se definir que há uma estação que

funcionará como servidora e concentrará a maior parte da demanda de tráfego, e que o volume não é uniforme entre as estações móveis;
- Após todo o período de experimentação com coleta dos pacotes no *switches* deve-se processar, de modo offline, esses pacotes por meio da geração de estruturas de dados do tipo *count-min* com contagem de Bytes e de pacotes:
  - Isso implica que para cada *switch* serão gerados 2 *sketches*, 1 com a contagem de Bytes e outro com a contagem de Pacotes;
  - O objetivo do uso desses contadores é uma representação estimada da matriz de tráfego entre os *switches*;
  - Sugere-se o uso da chave de *hashing* <srcIP, dstIP> para indexação das estruturas de dados. Outras alternativas podem ser exploradas (portas, endereços MAC e etc…);
- Sugere-se estabelecer um período fixo de amostragem para geração dos *sketches*, por exemplo, 1 *sketch* a cada 1 s. Outros valores podem ser usados;
- Posteriormente, deve-se gerar tráfego e/ou movimentação diferente do padrão (anômalo) estabelecido pelo grupo e coletar os pacotes para posterior processamento offline dos *sketches*;
- Com os *sketches* gerados a partir do tráfego anômalo, deve-se fazer uso de técnicas de Aprendizado de Máquina para avaliar e tentar detectar essas anomalias e avaliar os resultados. Sugere-se uso do K-Means, pelo menos como *baseline* de avaliação.

Bibliografia:

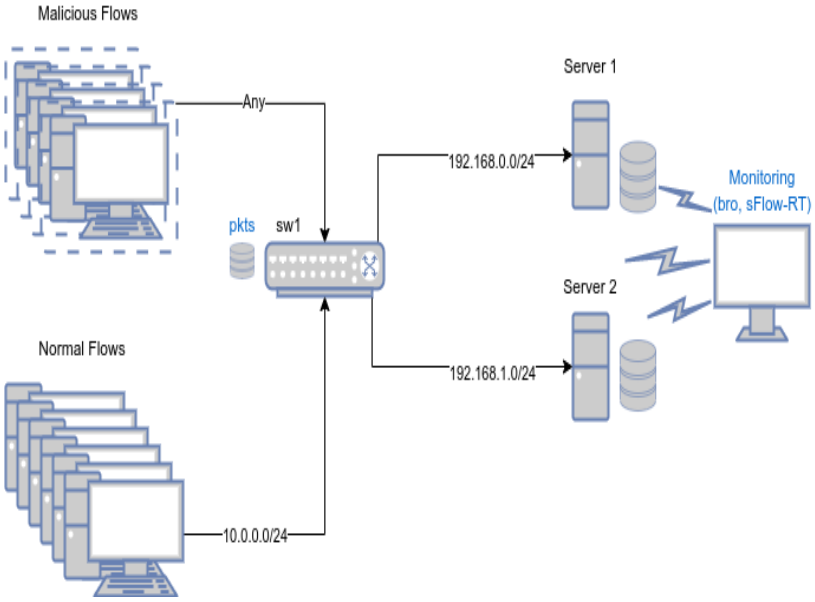
- Martins, R.F.T., da Silva Villaça, R. & Verdi, F.L. BitMatrix: A Multipurpose Sketch for Monitoring of Multi-tenant Networks. J Netw Syst Manage 28, 1745–1774 (2020).
- Leonardo Khoury Picoli. Caracterização e Detecção de Ataques de Negação de Serviço Usando Estruturas de Dados Probabilísticas e Aprendizado de Máquina. PG Engenharia de Computação, UFES.
- count-min sketch & its applications
- Silveira, Fernando Ávila Fossi. Minimização da Quantidade de Observadores Necessários para a Geração de Matrizes de Tráfego. Dissertação de Mestrado em Informática, PPGI/UFES.
- Aditya Chatterjee, Ethan Booker. Probabilistic Data Structures (Advanced Data Structures & Algorithms Book.
- Providing Mobility with Mininet-WiFi

Trabalho III – Detecção e Prevenção de Ataques de Intrusão

Integrantes:

- Vitor Fontana Zanotelli;
- Wagner Porto Ferreira;
- Israel dos Santos Candeias;
- Willen Borges Coelho (ouvinte).

Sugestão:



Requisitos:

- SW1: deve-se coletar as estatísticas de tráfego e todos os pacotes que trafegam nestes dispositivos para posterior processamento. Sugere-se o uso das ferramentas *wireshark* ou *tcpdump* para coleta de pacotes e do sFlow-RT para o monitoramento dos fluxos. Outras ferramentas, como *bro*[2] 9(ou similar) podem ser usadas nesta tarefa);
- Servers 1 e 2 devem estar em redes IP diferentes dos atacantes e dos clientes normais. Além disso, devem possuir sistema de monitoramento de métricas de sistema operacional e executar algum(ns) serviços de rede, por exemplo, HTTP;

2 https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1089.1520&rep=rep1&type=pdf

- Malicious Hosts: devem executar diversas sequências de *portscan* (tipo *nmap*) em instantes e origens distintas. Nenhum *malware* é necessário;
- Normal Flows: devem executar diversas sequências de *download* de páginas e arquivos web a partir de ambos os servidores, em instantes e origens diferentes.

Metodologia:

- Reproduzir, parcialmente, o artigo apresentado por (M. E SILVA, Gabriel Lucas F., 2022) usando *datasets* da base CTU-13. Neste contexto é extremamente importante compreender exatamente quais são as características do tráfego no *dataset* usadas como entrada para o treinamento da rede de detecção;
- Reproduzir ao menos 1 cenário escolhido no dataset CTU-13 a partir do reencaminhamento dos pacotes no Mininet;
  - Importante: não usar as redes *botnets* nesta etapa. Use os pacotes disponíveis no *dataset* e ferramentas de reprodução de tráfego para geração de outras estatísticas de monitoramento: sFlow-RT e avaliação do sistema operacional nos servidores.
- Implemente um cenário fictício para simular a preparação de um ataque por meio da existência prévia de uma sequência de verificação de portas (*portscan*) misturados a tráfego normal. Durante todo o experimento, colete pacotes e métricas de sistema operacional no servidor;
  - Importante: NÃO use a informação temporal e de IP de origem para treinamento dos algoritmos de detecção da varreção de portas. Pode-se usar essas informações para auxiliar na geração dos *labels* de aprendizado supervisionado, mas nunca para o treinamento;
  - Considere a possibilidade de usar o mesmo método usado no artigo de referência desta atividade (M. E SILVA, Gabriel Lucas F., 2022).

Referências:

- Zeek: An Open Source Network Security Monitoring Tool
- Vern Paxson. Bro: A System for Detecting Network Intruders in Real-Time. Usenix, 1998.
- M. E SILVA, Gabriel Lucas F.; NEIRA, Anderson Bergamini de; NOGUEIRA, Michele. Aprendizado Profundo para a Predição de Ataques de Negação de Serviço Distribuído. In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS (SBRC), 40. , 2022, Fortaleza. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2022. p. 475-488. ISSN 2177-9384.

- The CTU-13 Dataset. A Labeled Dataset with Botnet, Normal and Background traffic
- Detecção de Intrusão em Redes de Computadores. Cap 6.
- How to capture and replay network traffic on Linux
- Integration between Mininet-WiFi and sflow-rt